UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P O Box 1450
Alexandria, Virginia 22313-1450
www uspto gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/565,570	07/21/2006	Robert W. Morris	30004-004US1	2103

26161 7590 04/24/2008

FISH & RICHARDSON PC
P.O. BOX 1022
MINNEAPOLIS, MN 55440-1022

EXAMINER	
BORSETTI, GREG	
ART UNIT	**PAPER NUMBER**
4141	
MAIL DATE	**DELIVERY MODE**
04/24/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/565,570	MORRIS, ROBERT W.	
	Examiner	Art Unit	
	GREG A. BORSETTI	4141	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on _21 July 2006_.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1-18_ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1-18_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on _23 January 2006_ is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _2/12/2008,7/21/2006_.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

1. This action is in response to the preliminary amendment filed on 1/23/2006.

2. Claims 10,12 and 13 have been amended.

3. Claims 1-18 are pending.

Information Disclosure Statement

4. The Information Disclosure Statement (IDS) submitted on 7/21/2006 is not in compliance with the provisions of 37 CFR 1.97.

- US Patent (Desig. ID. AE - Charlesworth) has an incorrect publication date. It should be 3/29/2005, not 3/26/2005.
- NPL Document (Design. ID. AQ – Choi) was not included as NPL. A document titled "An Overview of the AT&T spoken document retrieval" was submitted with same author but since the titles do no match it is not understood which document is to be considered.
- NPL Document (Desig. ID. AR – Cooper) does not have a date.

5. The Information Disclosure Statement (IDS) submitted on 2/12/08is in compliance with the provisions of 37 CFR 1.97.

Drawings

6. The drawings filed on 1/23/2006 are accepted by the examiner.

Claim Rejections - 35 USC § 112

7. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claim 7 states "a number of recognition hypotheses" which is vague and indefinite. Clarification is needed.

Claim Rejections - 35 USC § 101

8. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-18 of the claimed invention are directed to non-statutory subject matter. Claims 1-18 lack a tangible output to make the invention provide useful, tangible, and concrete results which satisfy the conditions of 35 U.S.C. 101.

Claim Rejections - 35 USC § 102

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-3, 8-13, 15, and 17-18 are rejected under 35 U.S.C. 102(b) as being anticipated by Chou et al. (US Patent #5797123 hereinafter Chou).

As per claim 1, Chou discloses:

- **accepting query data from one or more spoken instance of a query**
- [Chou, column 4-5, lines 65-67, 1-9] discloses “In addition, however, the illustrative system of FIG. 1 goes even further to reduce such "false alarms." The system does not make a "final decision" as a result of these keyword (or key-phrase) matching and verification processes alone. Rather, a semantic analysis (i.e., sentence parsing) is performed based on combinations (i.e., sequences) of the verified keywords or key-phrases, resulting in sentence hypotheses which are then themselves verified with a separate verification process. In particular, this sentence hypothesis verification process is performed with a "partial input" comprising fewer subwords than are found in the entire utterance.” The input is an utterance which is a spoken instance of a query which is received.
- **processing the query data including determining a representation of the query that defines multiple sequences of subword units each representing the query**
- [Chou, column 5, lines 60-65] discloses “The subword model recognizer employed by key-phrase detector 11 uses lexicon 23 and subword models 22, which may have been trained based, for example, on a conventional minimum classification error (MCE) criterion, familiar to those skilled in the art.” The

subword model recognizer functionally works within the illustrative system in Fig 1.

- **locating putative instances of the query in input data from an audio signal**
- [Chou, column 5, lines 50-57] discloses "Specifically, the illustrative system of FIG. 1 includes key-phrase detector 11, key-phrase verifier 12, sentence hypothesizer 13 and sentence hypothesis verifier 14. In particular, key-phrase detector 11 comprises a subword-based speech recognizer adapted to recognize a set of key-phrases using a set of phrase sub-grammars (i.e., key-phrase grammars 21) which may advantageously be specific to the dialogue state." The key-phrase detector 11 locates instances of the query in input data from an audio signal.

As per claim 2, claim 1 is incorporated and Chou discloses:

- **processing the query data includes applying a speech recognition algorithm to the query data**
- [Chou, column 4, lines 30-42] discloses "(Subword-based speech recognition, familiar to those of ordinary skill in the art, involves the modeling and matching of individual word segments such as syllables, demisyllables or phonemes. A lexicon or dictionary is then provided to map each word in the vocabulary to one or more sequences of these word segments--i.e., the subwords. Thus, the model corresponding to a word effectively comprises a concatenation of the models for the subwords which compose that word, as specified by the lexicon.)

FIG. 1 shows a diagram of one illustrative system for performing speech recognition and understanding of a spoken utterance in accordance with an illustrative embodiment of the present invention." Fig .1 and the provided quotation describes how speech recognition is applied to the input query data.

As per claim 3, claim 1 is incorporated and Chou discloses:

- **subword units include linguistic units**
- [Chou, column 4, lines 23-33] discloses "In accordance with an illustrative embodiment of the present invention, a spoken dialogue recognition and understanding system is realized by recognizing the relevant portions of the utterance while not erroneously "recognizing" the irrelevant portions, (without, for example, using non-keyword large vocabulary knowledge) in a general framework of subword-based speech recognition. ***<u>(Subword-based speech recognition, familiar to those of ordinary skill in the art, involves the modeling and matching of individual word segments such as syllables, demisyllables or phonemes.</u>***" Chou discloses that subword units are used and furthermore states that phonemes make up the subword based speech recognition. It is well-known in the art that phonemes are the smallest forms of linguistic units, thus Chou teaches that the subword units used include linguistic units.

As per claim 8, claim 1 is incorporated and Chou discloses:

- **determining the representation of the query includes determining a network of the subword units**
- [Chou, column 6, lines 57-60] discloses “In particular, the key-phrase and filler-phrase grammars are compiled into networks, wherein key-phrases are recurrent and garbage models are embedded between key-phrase occurrences.” This goes in conjunction with the Fig. 1 showing how the subword-models 11 are utilized within the key-phrase detector 11 for its operation, thus there is a determination of a network of the subword units.

As per claim 9, claim 8 is incorporated and Chou, discloses:

- **multiple sequences of subword units correspond to different paths through the network**
- [Chou, column 6-7, lines 57-67, 1-5] discloses “In particular, the key-phrase and filler-phrase grammars are compiled into networks, wherein key-phrases are recurrent and garbage models are embedded between key-phrase occurrences. Note, however, that simple recurrence can result in ambiguity. For example, if any repetitions of the days of the month are allowed, it is not possible to distinguish between "twenty four" and "twenty"+"four." Therefore, additional constraints that inhibit impossible connections of key-phrases are incorporated as well. Therefore, the detection unit comprises a network of key-phrase sub-grammar automata with their permissible connections and/or iterations. Such automata can easily be extended to a stochastic language

model by estimating the connection weights. The use of such models achieves wider coverage with only modest complexity when compared with sentence-level grammars." Chou teaches that sub-grammars are used to differentiate between ambiguous terminology in network models. The sub-grammar automata are the subword units and they define different paths through the network to determine the meaning. [Chou, column 7, lines 5-15] teaches that Fig. 2 is reduced and does not show the sub-grammars that are used to define the words.

As per claim 10, claim 1 is incorporated and Chou discloses:

- **determining the representation of the query includes determining an n-best list of recognition results**
- [Chou, column 7, lines 47-57] discloses "When a hypothesis "popped" by the stack decoder has been tagged as a complete phrase to be output, the procedure extends the phrase by one additional word and aligns the phrase with the best extension. If this node is reached at the same time point by any of the previous hypotheses, then the current hypothesis is discarded after the detected phrase is output. Otherwise, the time point is marked for further search. Note that the detection procedure is quite efficient without redundant hypothesis extensions and ***produces the correct N-best key-phrase candidates in the order of their scores***."

As per claim 11, claim 10 is incorporated and Chou discloses:

- **each of the multiple sequences of subword units corresponds to a different one in the n-best list of recognition results**
- [Chou, column 7, lines 47-57] discloses that the stack decoder "***produces the correct N-best key-phrase candidates in the order of their scores.***" It does this without repetition, thus the subword units defining the phrases would inherently be unique as to the sequence of the subunits because if duplicates are detected, they are deleted.

As per claim 12, claim 1 is incorporated and Chou discloses:

- **accepting the query data includes accepting audio data representing the spoken utterances of the query spoken by a user, and processing the audio data to form the query data**
- [Chou, column 3, lines 49-52] discloses "First, a plurality of key-phrases are detected (i.e., recognized) based on a set of phrase sub-grammars which may, for example, be specific to the state of the dialogue. ***These key-phrases are then verified by assigning confidence measures thereto and comparing the confidence measures to a threshold, resulting in a set of verified key-phrase candidates***." The query data is developed from the processed input data. Furthermore, [Chou, column 4-5, lines 65-67, 1-9] discloses "In particular, this sentence hypothesis verification process is performed with a "partial input"

comprising fewer subwords than are found in the entire utterance." The input is an utterance which is a spoken instance of a query which is received.

As per claim 13, claim 1 is incorporated and Chou discloses:

- **accepting the query data includes accepting selection by a user of portions of stored data from a previously accepted audio signal, and processing the portions of the stored data to form the query data**
- [Chou, column 3, lines 47-52] discloses "Specifically, a "multiple pass" procedure is applied to a spoken utterance comprising a sequence of words (i.e., a sentence). First, a plurality of key-phrases are detected (i.e., recognized) based on a set of phrase sub-grammars which may, for example, be specific to the state of the dialogue. These key-phrases are then verified by assigning confidence measures thereto and comparing the confidence measures to a threshold, resulting in a set of verified key-phrase candidates." The key-phrase comparison would inherently have previously accepted audio signals in the defined key-phrases which are used for comparison. The resultant comparison forms the query data.

As per claim 15, claim 14 is incorporated and Chou discloses:

- **the first speech recognition algorithm produces data related to presence of the subword units at different times in the audio signal**

- [Chou, column 5, lines 60-65] discloses “The subword model recognizer employed by key-phrase detector 11 uses lexicon 23 and subword models 22, which may have been trained based, for example, on a conventional minimum classification error (MCE) criterion, familiar to those skilled in the art.” The subword model recognizer functionally works within the illustrative system in Fig 1. It would be inherent that there are subword units which are taken from different points in the audio signal to define the full range of the audio signal. For a full signal to be analyzed, there must be subword units for each definable subword unit meaning in the phrase, and the phrase would extend over a period of time, thus the subword units would as well. Furthermore, the speech Recognition algorithm would produce data from the subword units inherently, so it would also produce data related to presence of the subword units at different times in the audio signal.

Claims 17 and 18 are the software and hardware representations of the method as claimed in claim 1. Claims 17 and 18 are rejected under the same principles as claim 1 for having identical limitations. [Chou, column, lines] discloses "Illustrative embodiments of the present invention may comprise digital signal processor (DSP) hardware, ***read-only memory (ROM) for storing software performing the operations discussed above***, and random access memory (RAM) for storing results. ***Very large scale integration (VLSI) hardware embodiments***, as well as custom VLSI circuitry in combination with a general purpose processor or DSP circuit, may also be

provided." Chou provides software and hardware illustrative embodiments which anticipate claims 17 and 18.

Claim Rejections - 35 USC § 103

10. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 4-7, and 14 are rejected under 35 U.S.C. 102(b) as being taught by Chou et al. (US Patent #5797123 hereinafter Chou).

As per claim 4, claim 2 is incorporated and Chou teaches:

- **locating the putative instances includes applying a word spotting algorithm configured using the determined representation of the query**
- [Chou, column 5, lines 10-19] discloses "As pointed out above, the illustrative system of FIG. 1 advantageously uses key-phrases as the detection unit rather than using only keywords. Typical word spotting schemes as described above use small templates that can easily be triggered by local noise or confusing sounds. Using longer units of detection (i.e., key-phrases instead of just keywords) is advantageous because it tends to incorporate more distinctive information, resulting in more stable acoustic matching, both in the recognition

phase and in the verification phase." Chou discloses that word spotting algorithms are well known in the art and that a key-phrase detection unit is used rather than just a word spotting scheme for further accuracy. Thus, it would be obvious to someone of ordinary skill to use a word spotting scheme because it is well known in the art.

As per claim 5, claim 4 is incorporated and Chou teaches:

- **selecting parameter values of the speech recognition algorithm for application to the query data according to characteristics of the word spotting algorithm**
- [Chou, column 5, lines 27-49] discloses "In accordance with the illustrative embodiment of the present invention described herein, the detected key-phrases are advantageously tagged with conceptual information. In fact, the key-phrases may be defined so as to directly correspond with semantic slots in a semantic frame, such as, for example, a time and a place.... the top-down key-phrases recognized by the instant illustrative embodiment may easily be directly mapped into semantic representations. Thus, the detection of these key-phrases directly leads to a robust understanding of the utterance." The key-phrase detector, which has been shown above to be an obvious replace for the key-spotting detection, tags detected phrases with conceptual information for further consideration by the speech recognition algorithm.

As per claim 6, claim 5 is incorporated and Chou teaches:

- **selecting of the parameter values of the speech recognition algorithm includes optimizing said parameters according to an accuracy of the word spotting algorithm**
- [Chou, column 5, lines 60-67] discloses "The subword model recognizer employed by key-phrase detector 11 uses lexicon 23 and subword models 22, which may have been trained based, for example, on a conventional minimum classification error (MCE) criterion, familiar to those skilled in the art. The models themselves may, for example, comprise Hidden Markov Models (i.e., HMMs), also familiar to those skilled in the art." The parameters are optimized according to the phrase spotting, an obvious replacement for the word spotting, for the speech recognition algorithm.

As per claim 7, claim 5 is incorporated and Chou teaches:

- **selecting of the parameter values of the speech recognition algorithm includes selecting values for parameters including one or more of an insertion factor, a recognition search beam width, a recognition grammar factor, and a number of recognition hypotheses**
- [Chou, column 6, lines 35-57] discloses "Specifically, for each sub-task, key-phrase patterns are described as one or more deterministic finite state grammars, illustratively selected by key-phrase detector 11 from key-phrase grammars 21. These grammars may be manually derived directly from the task

specification, or, alternatively, they may be generated automatically or semi-automatically (i.e., with human assistance) from a small corpus, using conventional training procedures familiar to those skilled in the art." A recognition grammar factor is used.

As per claim 14, claim 13 is incorporated and Chou discloses:

- **prior to accepting the selection by the user, processing the previously accepted audio signal according to a first speech recognition algorithm to produce the stored data**
- [Chou, column 5, lines 60-67] discloses "The subword model recognizer employed by key-phrase detector 11 uses lexicon 23 and subword models 22, which may have been trained based, for example, on a conventional minimum classification error (MCE) criterion, familiar to those skilled in the art. The models themselves may, for example, comprise Hidden Markov Models (i.e., HMMs), also familiar to those skilled in the art." Furthermore, [Chou, column 6, lines 40-45] discloses "These grammars may be manually derived directly from the task specification, or, alternatively, they may be generated automatically ***or semi-automatically (i.e., with human assistance) from a small corpus***, using conventional training procedures familiar to those skilled in the art." It would be obvious to someone of ordinary skill in the art that semi-automatically trained grammar could be trained with phrasal analysis by the speech

recognition algorithm as would be well known to someone of ordinary skill in the art.

Claim 16 is rejected under 35 U.S.C. 102(b) as being anticipated by Chou et al. (US Patent #5797123 hereinafter Chou) in view of Thong et al. (US Pre-Grant Publication #20030110035 hereinafter Thong).

As per claim 16, claim 14 is incorporated and Chou fails to teach:

- applying a second speech recognition algorithm to the query data

However, in analogous art, Thong teaches the above limitation,

- [Thong, Fig. 2] discloses the use of two separate algorithms applied to the query data. The first being the word comparison and the second being the subword comparison analogous to Chou.
- Thong and Chou are analogous are because both deal with word-spotting and subword detection for speech recognition. It would be obvious to someone of ordinary skill in the art to combine Thong with the Chou device because "The method of the present invention allows modeling user input so as to take into account the acoustic inaccuracy by returning the most likely answers to the user." The Thong addition would benefit the Chou device by taking acoustic inflection into consideration in its speech recognition method.

Conclusion

11. Refer to PTO-892, Notice of References Cited for a listing of analogous art.

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to GREG A. BORSETTI whose telephone number is (571)270-3885. The examiner can normally be reached on Monday - Thursday (8am - 5pm Eastern Time).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Chameli Das can be reached on 571-272-3696. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Greg A. Borsetti/
Examiner, Art Unit 4141

/CHAMELI C. DAS/
Supervisory Patent Examiner, Art Unit 4141
Dated: 4/23/08